

**IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF PENNSYLVANIA**

SHARYN STAGI and WINIFRED LADD	:	CIVIL ACTION
Individually, and on behalf of	:	
all others similarly situated	:	
	:	
Plaintiffs,	:	
	:	NO. 03-CV-5702
v.	:	
	:	
NATIONAL RAILROAD PASSENGER	:	
CORPORATION (t/d/b/a AMTRAK)	:	
	:	
Defendant.	:	

MEMORANDUM

Anita B. Brody, J.

August _____, 2009

Plaintiffs Sharyn Stagi and Winifred Ladd bring this civil action against the National Railroad Passenger Corporation (“Amtrak”), asserting that a company policy that requires all union employees to have one year of service in their current position before they will be considered for promotion, has a disparate impact on female union employees in violation of Title VII of the Civil Rights Act of 1964, 42 U.S.C. 2000e, and the Equal Protection component of the Due Process Clause of the Fifth Amendment. Because the plaintiffs’ evidence of disparate impact lacks both statistical and practical significance, the plaintiffs have failed to make out a *prima facie* case of discrimination under Title VII.

I. Background

This case concerns an allegedly discriminatory employment policy of the defendant,

Amtrak. Amtrak's workforce is broadly divided into two groups of employees: union employees whose transfer and promotion rights within their particular unions are governed by collective bargaining agreements, and management employees, whose transfer and promotion are controlled through Amtrak's own employment policies and procedures. Plaintiffs Sharyn Stagi and Winifred Ladd are long-service employees of the defendant, and have held both union and non-union positions. Stagi began her career with Amtrak in 1973 as an entry-level union employee, in the capacity of reservation and information clerk. In the early 1990s she was promoted to Inventory Control Planner, and again later to System Analyst - a non-union, management position. Ladd had a similar career at Amtrak, also starting as a reservation and information clerk in 1973, and promoted within Amtrak's union ranks until she secured the supervisory management position of Operation Support Speciality in 1986. In April 2002, both Stagi and Ladd were laid off from their management level positions as a result of a corporate-wide management restructuring effort. However, because both plaintiffs had previously held union positions and had retained their union membership, they were entitled to "bump-down," or place bids to take non-management union jobs, albeit with lower pay and benefits. Both plaintiffs took advantage of this policy and secured union positions with Amtrak.¹

Within a year of being laid off, Amtrak posted vacancies for management positions. Stagi and Ladd applied to these positions and their applications were rejected.² Stagi and Ladd

¹ The plaintiffs contend that a majority of management-level employees affected by the 2002 layoffs were female, due to their lower seniority and pay grade. A majority of laid-off management-level employees who bumped-down to union positions following the 2002 layoffs also were female.

² In Stagi's case, one of the positions she applied to, and was rejected from, was a management position she had previously held at Amtrak.

were notified that they could not be considered for these management positions because of an Amtrak policy known as PERS-4, which requires all union employees to have worked in their current position for at least one year before being considered for a posted management job (the “one-year rule” or the “Policy”). The Policy states in full:

A non-agreement covered employee may not apply for a posted non-agreement covered position if he or she has not been in his or her current position for at least one year.³

An agreement covered employee may not apply for a posted non-agreement covered position unless he or she has been in his or her current union [sic] for one year.⁴

However, if these restrictions create a hardship for Amtrak, the employee’s supervisor, with the approval of the Human Resources Department, may grant an exception to this rule. (Am. Compl. Ex. A.)

The plaintiffs contend that the portion of the Policy that applies to employees seeking to move from agreement covered (union) jobs to non-agreement covered (management) jobs has a

³ The term “non-agreement covered” refers to non-union, or management positions. Thus, this portion of the Policy applies only to lateral job movements within the management ranks. In their complaint, the plaintiffs challenge only the portion of the Policy that applies to job movements from union (agreement covered) positions to management (non-agreement covered) positions. Amtrak argues that because plaintiffs attack only union to management movements, their analysis does not accurately reflect the way the one-year rule is actually implemented at Amtrak. Because the court finds that the plaintiffs have failed to make out a *prima facie* case on the evidence presented with respect to union-to-management promotions, it is unnecessary to reach this argument.

⁴ Although the way the rule is written appears to prevent consideration of agreement-covered employees based on time-in-current-*union*, since at least 1999 or 2000, the Policy has been applied consistently to consider time-in-*position*, not time-in-union. (emphasis added) (Green Dep. 170:1-12, Sept. 30, 2004.) The language of the policy was changed in 2004 (after the commencement of this litigation) and currently reads: “An agreement covered (union) employee may not apply for a posted non-agreement covered (management) position unless he or she has been in his or her current agreement covered (union) position for one year.” (Pl.’s Mem. Ex. 4.)

discriminatory impact on female employees because it has the effect of denying management opportunities to a disproportionate number of female employees, relative to male employees, in violation of Title VII of the Civil Rights Act of 1964 and the Equal Protection Clause of the Fourteenth Amendment, as incorporated into the Fifth Amendment.

Plaintiffs filed their complaint in this action on October 14, 2003. The complaint was amended on May 25, 2004 (Doc. #9). On December 30, 2005, the Court denied Amtrak's motion for judgment on the pleadings pursuant to Fed. R. Civ. P. 12(c). Discovery proceeded and on April 4, 2007, the court held a discovery conference during which the court extended the close of fact discovery, and indicated that ruling on summary judgment would precede ruling on class certification (Doc. #58; *see also* Doc. #61). On February 29, 2008, the plaintiffs moved for class certification pursuant to Fed. R. Civ. P. 23 (Doc. #70). On April 21, 2008, Amtrak moved for Summary Judgment (Doc. #77). The Court ordered oral argument regarding both class certification and summary judgment, together, on September 8, 2008. Pursuant to the parties' request, oral argument on all motions was postponed until after summary judgment briefing was completed. (*See* Joint letter from counsel, dated September 4, 2008). On July 21, 2009, the Court heard oral argument and received testimony from the expert witnesses in anticipation of ruling on class certification.⁵ Both class certification and summary judgment motions are now

⁵ The parties arguments for and against class certification, particularly with respect to whether the putative class met the Rule 23 requirements of commonality and typicality, hinged on the validity of the plaintiffs' expert report as evidence of classwide discrimination - a dispute a district court is required to resolve under *In re Hydrogen Peroxide*, 552 F.3d 305, 323 (3d Cir. 2008) (holding that the court must resolve all factual or legal disputes relevant to class certification, and that expert opinion relating to a Rule 23 requirement is no exception). To this end, the argument and testimony presented to the court on July 21, 2009 was focused on explanation of the plaintiffs' expert report, and the defendant's rebuttal report as it related to class certification. As further explained in this summary judgment opinion, the plaintiffs' expert

pending before the court.

II. Summary Judgment Standard

Summary judgment is appropriate “if the pleadings, the discovery and disclosure materials on file, and any affidavits show that there is no genuine issue as to any material fact and that the movant is entitled to judgment as a matter of law.” Fed. R. Civ. P. 56(c); *Kornegay v. Cottingham*, 120 F.3d 392, 395 (3d Cir. 1997). A factual dispute is “genuine” if the evidence would permit a reasonable jury to find for the non-moving party. *Anderson v. Liberty Lobby Inc.*, 477 U.S. 242, 248 (1986). The party moving for summary judgment bears the initial burden of demonstrating that there are no facts supporting the nonmoving party’s legal position. *Celotex Corp. v. Catrett*, 477 U.S. 317, 323 (1986). Once the moving party carries this initial burden, the nonmoving party must set forth specific facts showing that there is a genuine issue for trial. Fed. R. Civ. P. 56(e); *see also Matsushita Elec. Indus. Co., Ltd. v. Zenith Radio Corp.*, 475 U.S. 574, 587 (1986). The non-moving party “cannot rely merely upon bare assertions, conclusory allegations or suspicions to support its claim.” *Fireman’s Ins. Co. v. DeFresne*, 676 F.2d 965, 969 (3d Cir. 1982). Rather, the party opposing summary judgment must go beyond the pleadings

report is insufficient evidence that all female employees of Amtrak have been disparately impacted in seeking and obtaining management positions as a result of the one-year rule, a finding that would have defeated class certification as well. *See e.g., Ross v. Nikko Securities Co. Intern., Inc.*, 133 F.R.D. 96 (S.D.N.Y. 1990) (denying class certification where the plaintiffs’ could not establish “an aggrieved class of female employees because the statistics...failed to offer ‘the relevant comparisons of similarly situated female and male employees (i.e., female and males with the same qualifications and experience)’”) (quoting *Sheehan v. Purolator, Inc.*, 103 F.R.D. 641, 649 (E.D.N.Y. 1984), *aff’d* 839 F.2d 99 (2d Cir. 1988)); *compare Hnot v. Willis Group Holdings Ltd.*, 228 F.R.D. 476 (S.D.N.Y. 2005) (finding commonality after reviewing expert report and determining that the analysis was rigorous enough to establish common questions of fact regarding the discriminatory implementation of the challenged policies).

and present evidence, through affidavits, depositions, or admissions on file, to show that there is a genuine issue for trial. *Celotex*, 477 U.S. at 324. The court must draw all reasonable inferences in the non-moving party's favor. *Matsushita*, 475 U.S. at 587. In a disparate impact employment discrimination case, summary judgment is warranted if the plaintiff fails to make out a *prima facie* case of discrimination. See *Foxworth v. Pennsylvania State Police*, 228 Fed. App'x. 151, 155-56 (3d Cir. 2007) (affirming district court's grant of summary judgment where plaintiff's statistical evidence failed to make out a *prima facie* case of disparate impact).

III. Establishing a Claim of Disparate Impact Gender Discrimination Under Title VII

Title VII of the Civil Rights Act of 1964, 42 U.S.C. § 2000e, prohibits employers from discriminating against any individual with respect to hiring, or the terms and conditions of employment, and from limiting, segregating, or classifying “employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual’s race, color, religion, sex, or national origin.”⁶ In *Griggs v. Duke Power Co.*, the Supreme Court construed Title VII to proscribe both overt instances of discrimination, and also “practices that are fair in form, but discriminatory in operation.” 401 U.S. 424, 431 (1971). Thus, an employment practice that is facially neutral may, in certain cases, be deemed violative

⁶ Title 42 U.S.C. § 2000e-2(a) provides in full: “It shall be an unlawful employment practice for an employer- (1) to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to his compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin; or (2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual's race, color, religion, sex, or national origin.” 42 U.S.C. § 2000e-2(a).

of Title VII if it has a disproportionate effect on a protected group. This basis for liability under Title VII is known as “disparate impact” discrimination.

Unlike discrimination cases addressing “disparate treatment,” disparate impact cases do not require proof of the employer’s subjective intent to discriminate. *Griggs*, 401 U.S. at 432; *Wards Cove Packing Co., Inc. v. Atonio*, 490 U.S. 642, 645 (1989). A *prima facie* case of disparate impact discrimination requires that the plaintiff first identify “the specific employment practice that is challenged.” *Watson v. Fort Worth Bank and Trust*, 487 U.S. 977, 994 (1988). Second, the plaintiff must show “causation”; in other words, that the employment practice “causes a disparate impact on the basis of race, color, religion, sex, or national origin.” 42 U.S.C. § 2000e-2(k)(1)(A)(i). To show causation, the plaintiff must offer “statistical evidence of a kind and degree sufficient to show that the practice in question has caused the exclusion of applicants for jobs or promotions because of their membership in a protected group.” *Watson*, 487 U.S. at 994; *see also Foxworth*, 228 Fed. App’x at 156. This means that statistical disparities must be “sufficiently substantial” such that they raise “an inference of causation.” *Watson*, 487 U.S. at 994-95.

There is no “rigid mathematical formula” that satisfies the “sufficiently substantial” standard in the disparate impact analysis, *id.*, however, the Equal Employment Opportunity Commission (“EEOC”) has provided some guidance in the EEOC’s Uniform Guidelines on Employee Selection Procedures, 29 CFR § 1607.4 (D) (1987). While, the EEOC Uniform Guidelines are not binding on courts, the Supreme Court has indicated that the guidance of this administrative body should be considered with “great deference,” and no consensus has developed around any alternative standard. *Griggs*, 401 U.S. at 433-34; *Watson*, 487 U.S. at 995

n.3. According to the Guidelines, evidence that a selection rate for any group is “less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest [selection] rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by the Federal enforcement agencies as evidence of adverse impact.” 29 CFR § 1607.4 (D) (1987). This standard, known as the “Four-Fifths Rule,” is not intended to be an absolute requirement: “Smaller differences in selection rate may nevertheless constitute adverse impact, where they are significant in both statistical and practical terms or where a user’s actions have discouraged applicants.” *Id.* Notwithstanding the EEOC’s “rule of thumb,” courts have recognized that in the “complex area of employment discrimination,” statistics ““come in infinite variety and ... their usefulness depends on all of the surrounding facts and circumstances.”” *Watson*, 487 U.S. at 995 n.3 (quoting *Teamsters v. United States*, 431 U.S. 324, 340 (1977)). For this reason, when faced with decisions as to whether the statistical evidence presented demonstrates adverse impact, courts generally evaluate the evidence on a case-by-case basis. *Id.*

If a plaintiff succeeds in making an initial showing of a specific employment practice and a statistical disparity that raises an inference of causation, the burden shifts to the defendant to show that the challenged practice is “job related for the position in question and consistent with business necessity.” 42 U.S.C. § 2000e-2(k)(1)(A)(i); *see also Foxworth*, 228 Fed. App’x at 156. In this phase, “the employer carries the burden of producing evidence of a business justification for his employment practice.” *Wards Cove*, 490 U.S. at 659. However, the ultimate burden of proving discrimination against a protected group because of a specific employment practice remains with the plaintiff at all times. *Id.* Thus, if the employer meets the burden of proving that

the challenged employment practice is “job related,” it falls to the plaintiff to “show that other tests or selection devices, without a similarly undesirable [discriminatory] effect, would also serve the employer’s legitimate interest in efficient and trustworthy workmanship.” *Albemarle Paper Co. v. Moody*, 422 U.S. 405, 425 (1975) (internal quotations omitted); *see also* 42 U.S.C. § 2000e-2(k)(1)(A)(ii).

IV. Discussion

Plaintiffs have satisfied the first part of their *prima facie* case of disparate impact by identifying PERS-4, or Amtrak’s one-year rule, as the specific employment practice being challenged. To satisfy the second part of the test, the plaintiffs must produce statistical evidence of a kind and degree sufficient to show that the one-year requirement has caused the disproportionate exclusion of women from consideration for promotions because of their membership in a protected group.

1. Plaintiff’s Statistical Evidence of Disparate Impact

In support of their *prima facie* case, the plaintiffs submit the expert report of Dr. Mark Killingsworth, who concludes from his analysis that Amtrak’s female union employees are blocked from consideration for promotion at a higher rate than men, as a result of the one-year rule.⁷ In performing his analysis, Dr. Killingsworth used employee data provided by Amtrak that included each employee’s gender, the date(s) on which they started each position they have held since the beginning of 2001, the position title and position “ID number,” and whether the job is union or management. Dr. Killingsworth also relied on a document provided by Amtrak’s Senior

⁷ There is no dispute that the parties’ experts are qualified to provide the opinions expressed in this case.

Director of Human Resource Operations, Michael Ramirez, that identifies position titles regarded by Amtrak as equivalent to other position titles.⁸

Using this employee data, Dr. Killingsworth first identified all incidences between March 8, 2002, and June 30, 2007, where a union employee moved from a union position to a management position. Each of these incidences is referred to as a “job fill.” Dr. Killingsworth’s report does not indicate the total number of “job fills” that occurred during this period, but defendant’s expert, Dr. Griffin, calculated that Dr. Killingsworth analyzed 716 “job fills” in his analysis.⁹ (Griffin Rep. 2.)

These approximately 700 “job fills” represent actual promotions to management resulting from actual hiring decisions at Amtrak during the period in question. Typically, the next step for a disparate impact analysis would be to compare the group of individuals who had actually applied for those promotions against those who had actually been hired. *Wards Cove*, 490 U.S. at 650 (holding that the proper comparison in disparate impact cases is usually between the minority composition “of the qualified persons in the labor market and the persons holding at-issue jobs”). An exception to this general proposition is warranted in so-called “entrance requirement” cases, because “persons who lack the challenged requirement [may] self-select themselves out of the pool of applicants.” *Moore v. Hughes Helicopters, Inc.*, 708 F.2d 475, 482 (9th Cir. 1983); *see also Dothard v. Rawlinson*, 433 U.S. 321, 330-31 (1977). This case

⁸ The parties have also stipulated to the accuracy of the underlying dataset (the employment histories of Amtrak’s employees) produced by Amtrak and used by both experts.

⁹ Although Dr. Killingsworth analysis is premised upon 716 “job fills,” Dr. Griffin was unable to locate 20 promotion events in the data set, leading Dr. Griffin to conclude that Dr. Killingsworth’s analysis is actually keyed to 696 promotion events. Dr. Killingsworth does not rebut or clarify these conclusions in his subsequent affidavits.

presents just such a problem, as the challenged Policy - the one-year experience requirement - may discourage otherwise qualified individuals with less than one year of experience from applying. Thus, the pool of actual applicants is likely to be under-representative of the group whom the Policy is truly affecting. In such cases, it is proper to establish disparate impact “through reference to a reasonable proxy for the pool of individuals actually affected by the alleged discrimination.” *Moore*, 708 F.2d at 482. The choice for statisticians faced with this dilemma is “usually between general population statistics and the statistics of a relevant labor market.” *Id.*

In this case, Dr. Killingsworth sought to approximate the group of employees affected by the Policy by identifying other employees who were similarly situated to an individual who was actually promoted to a particular management position on a particular date. In other words, if an individual was promoted from his or her union position to management, Dr. Killingsworth assumed that all other individuals holding that union position at that time were presumptively qualified for, and interested in, consideration for that promotion, and therefore subject to the one-year rule.

To identify these individuals, for each promotion, or “job fill” at issue, Dr. Killingsworth identified the union job title (or equivalent) that each employee held just prior to their promotion. This prior job title (or equivalent) was then identified as a “feeder job” for that “job fill” on that date. For example, if an employee moved to a management position from the union position of “Ticket/Accounting Clerk” on January 2, 2009, then the job of “Ticket/Accounting Clerk” would be classified as a “feeder job” for that management position as of that date. Next, all employees holding the “feeder job” title on the date immediately preceding each “job fill” were pooled

together into one group (the “Feeder Pool”). Using the same example, if employee A was promoted from the union position Ticket/Accounting Clerk to a management position on January 2, 2009, all employees who were Ticket/Accounting Clerks on January 1, 2009 would be considered part of the Feeder Pool for that management position on that date.

Dr. Killingsworth does not consider whether other job titles, outside of those job titles from which a promotion occurred, could have acted as a “feeder job” for any given promotion to management. Only employees who held a job that acted as a “feeder job” at some point between March 8, 2002, and June 30, 2007, are considered to be in the relevant labor market for Dr. Killingsworth’s original analysis. Although Amtrak argues at different points that these pools might be over or under-inclusive, Amtrak ultimately agrees that Dr. Killingsworth’s original feeder job model is consistent with Amtrak’s own hiring practices: While no rule prohibits union employees from applying for any position, as a practical matter, only applicants meeting minimum qualifications for the management job in question are given serious consideration. (*See* Def.’s Reply In Support of Mot. Summ. J. 14 n.10; *see also* Ramirez Dep., 204-206, Sept. 6, 2007.) In the absence of explicit measures of qualifications and job interest,¹⁰ Dr. Killingsworth assumed that information about the position held prior to promotion could reasonably serve as an indicator of qualifications and job interest. Based on the information provided to Dr. Killingsworth by Amtrak, plaintiffs’ method is a reasonable one.

Dr. Killingsworth repeated his process of creating a Feeder Pool for each movement from a union position to a management position that occurred between March 8, 2002 and June 30,

¹⁰ Although the record is clear that Amtrak did expect applicants to have minimum qualifications, Amtrak failed to explicitly identify them.

2007, resulting in approximately 716 separate Feeder Pools. Next, Dr. Killingsworth aggregated the Feeder Pools into one giant pool (the “Aggregated Feeder Pool”) and analyzed the degree to which the Policy disqualified women in the Aggregated Feeder Pool relative to men.

The total number of female employees in the Aggregated Feeder Pool was 29,919. Of those, 8,310 (or 27.77 percent), had less than one year of experience in their current position (“within-position service”). (Killingsworth Rep. Table 1.) The total number of male employees in the Aggregated Feeder Pool was 102,156, of which 25,896 (or 25.35 percent) had less than one year of within-position service. Thus, 27.77 percent of females and 23.35 percent of men in the Aggregated Feeder Pool did not meet the one-year requirement on a date that a promotion was made from their job rank, and were therefore blocked from consideration for promotion on that date.¹¹

The difference between men and women who were ineligible for promotion in the Aggregated Feeder Pool was 2.42 percent. Using a conventional chi-square test, Dr. Killingsworth found that this difference of 2.42 percentage points was the equivalent of 8.42 “standard deviation units.” (Killingsworth Rep. Table 1.) These results were confirmed by a probit analysis using “robust standard errors” to correct for the fact that the same individual might appear in more than one pool (and may therefore appear more than once in the data set).¹²

¹¹ It is important to remember that these numbers do not represent individuals who actually applied for and were rejected from promotion, but represent individuals who, because they held a “feeder job,” were presumed minimally qualified for promotion at some point over the time period at issue (March 8, 2002 through June 30, 2007), and therefore were presumptively affected by the one-year rule.

¹² Because union employees have the ability to change jobs laterally within the union ranks at any time, employee A might have been a Ticket/Accounting Clerk on January 1, 2009, and counted in that Feeder Pool if a “job fill” occurred from that position on that date, but if he

(*Id.* at Table 3.) The results of the corrected probit analysis yielded a standard deviation of 3.855.

(*Id.*) Any number of standard deviations at least equal to 2 is considered statistically significant at conventional test levels, or in other words, unlikely to have occurred as a result of chance alone.¹³ See *Castaneda v. Partida*, 430 U.S. 482, 496 n.17 (1977). Based on these analyses, Dr. Killingsworth concluded that the one-year rule has disproportionately blocked women, relative to men, who might otherwise have been considered for promotion from union positions into management positions. (Killingsworth Rep. 5.)

As discussed, *supra*, mere demonstration of a statistical disparity is not sufficient to carry the *prima facie* burden in a disparate impact case. Rather, the statistical evidence must be “sufficiently substantial” such that it raises an inference of causation. *Watson*, 487 U.S. at 994-95. It is undisputed that the results of Dr. Killingsworth’s analysis would not indicate discrimination under the EEOC’s Four-Fifths Rule. In fact, the adverse impact ratio of Dr.

or she then switched to another “feeder job” from which a “job fill” occurred within the relevant period, he or she would be counted in that Feeder Pool as well. It is unclear from Dr. Killingsworth’s report how he accounted for an individual who appeared in the Aggregated Feeder Pool as having more than one year of in-position service at one time, and less than one year of in-position service at another time. For example, if employee A who was a Ticket/Accounting Clerk on January 1, 2009 had more than one year of in-position service when a “job fill” occurred from that position, employee A would be counted as eligible for promotion on that date. If employee A then made a lateral move to another union job that became a feeder job less than one year later, that same individual would be deemed ineligible within that Feeder Pool. Because Dr. Killingsworth’s analysis aggregates data from multiple positions over a four year period, the same individual may be in four or less groups - and eligible for promotion each time - or in dozens of groups, ineligible each time. Dr. Killingsworth does not explain how he controls for individuals with a potentially shifting ineligibility status. As explained, *infra*, failing to control for differences in ineligibility rates among employees over time and position undermines the utility of his report as evidence of disparate impact.

¹³ The “p-value” or “significance level” is a calculation as to the probability that a difference at least as large as the one actually obtained could have occurred as a result of chance alone. In this case, 3.8 standard deviations resulted in a p-value of <0.001.

Killingsworth's analysis is well above 80% at 96.8% (100% meaning identical eligibility rates by gender) (Griffin Rep. 1-2.) Thus, in order to determine whether the plaintiffs' evidence is "sufficiently substantial," the court must consider the numerical disparities in terms of both its statistical significance and practical significance.

2. Defendant's Rebuttal: Plaintiffs' Evidence is Not Sufficiently Substantial

Amtrak argues that the plaintiffs' statistical evidence is neither statistically significant nor practically significant, as a matter of law. With respect to statistical significance, Amtrak argues that Dr. Killingsworth's methodology of aggregating the Feeder Pools before performing his statistical analysis is fundamentally flawed, and that had Dr. Killingsworth remained faithful to his concept of Feeder Pools throughout his analysis, he would have found *no* statistically significant likelihood that male employees would be promoted over female employees. In the alternative, Amtrak argues that while Dr. Killingsworth's results may demonstrate "statistical significance," they do not reflect the practical significance required to raise an inference of discrimination in a disparate impact case. In support of these arguments, Amtrak submits the report of their retained expert, Dr. David W. Griffin. *See Watson*, 487 U.S. at 996 ("If the employer discerns fallacies or deficiencies in the data offered by the plaintiff, he is free to adduce countervailing evidence of his own.") (quoting *Dothard*, 433 U.S. 321, 331 (1977)).

a. Statistical Significance

Amtrak argues that Dr. Killingsworth's report is internally inconsistent and unreliable, and therefore insufficient to prove a *prime facie* case of discrimination. The inconsistency stems from Dr. Killingsworth having used one methodology (stratification) to compile a database of separate eligibility pools for each of the approximately 700 management vacancies at issue, and

then changing that methodology to perform the statistical analysis on the entire data set as a whole (aggregation). Amtrak argues that changing the methodology destroys the credibility of the report because it has the effect of treating every union position as a feeder for every management position - an unrealistic model of Amtrak's historical union-to-management promotion patterns. Dr. Griffin's analysis also demonstrates that if Dr. Killingsworth had remained faithful to his Feeder Pool model throughout his analysis, he would have found no evidence of disparate impact.

1. Dr. Griffin's Stratification Model

In order to highlight the flaws in Dr. Killingsworth's analysis, Dr. Griffin starts with the same set of "job fills" and Feeder Pools that Dr. Killingsworth constructed from Amtrak's employee data. However, instead of analyzing the Feeder Pools together *en masse*, Dr. Griffin analyzes each "job fill" on a vacancy-by-vacancy basis and asks whether the number of ineligible women in each pool was greater or less than what one would "expect," given the rate of ineligibility for the pool as a whole.¹⁴ Sometimes these calculations yielded a result adverse to females (meaning there were greater than expected ineligible females), sometimes the result was adverse to males (meaning there were greater than expected eligible females), and sometimes the result indicated no gender difference whatsoever (gender parity). What Dr. Griffin finds in

¹⁴ This calculation is made in three steps. First, Dr. Griffin determines the percentage of ineligible men and women in the Feeder Pool (i.e., if 30 out of 300 people are blocked, the total ineligibility rate would be 10%). Second, Dr. Griffin multiplies that percentage by the total number of women in the pool to devise the number of "expected" ineligible, given the makeup of that particular pool (i.e., if there were 200 women in the pool, multiplied by 10%, you would "expect" 20 women in the pool to be ineligible given the rate for the pool as a whole). Third, Dr. Griffin compared the "expected" number of ineligible women with the actual number of ineligible women in the pool, to assess whether there was a shortfall or a surplus of ineligible women in that particular pool, relative to what was "expected."

stratifying the data this way is that, depending on the makeup of the employees in each Feeder Pool in terms of gender and length of within-position service, the ineligibility rate for that group changes, and a different pattern of impact emerges - one where females are disadvantaged vis-à-vis men by operation of the one-year rule for some “job fills,” but not others. Dr. Griffin provides an illustration of this phenomenon in the case of a promotion to Trainmaster that occurred on October 16, 2006. For that particular promotion, because of the gender composition of the Feeder Pool and the rate of ineligibility for the group as a whole, males in that Feeder Pool were actually impacted by the one-year rule at a greater rate than women.¹⁵ Thus, in Dr. Griffin’s calculation the women presumed eligible for promotion to Trainmaster on October 16, 2006, are not disadvantaged by the one-year rule, whereas women qualified for promotion to a different position on that date may be severely impacted by the one-year rule because they face a different configuration of peer applicants. When looked at as individual promotion events, the impact of the one-year rule on female union employees becomes a metric that shifts according to who is in the Feeder Pool at the time of the hiring event. Simply stated: on a promotion-by-promotion basis, it is not true that women are always disadvantaged relative to men. (Griffin Rep. 6.)

To analyze whether women might still be disadvantaged relative to men overall, Dr. Griffin summed the surpluses and shortfalls of ineligible females across approximately 600 “job fills.”¹⁶ The results yielded a net surplus of 6.2 ineligible females, which translates to six fewer

¹⁵ Dr. Griffin used the Trainmaster example as part of his illustration of “Simpson’s Paradox,” wherein an examination of two promotions in isolation showed men suffering greater impact than women in each case, even though the overall rate of ineligibility for both positions combined actually favored men and disfavored women.

¹⁶ Dr. Griffin collapsed Dr. Killingsworth’s data set into 637 mutually exclusive promotion events, based on (1) the management position being filled; the “fill” date; and (3) the

promotion-eligible females than what gender parity within every pool would lead one to expect. Six fewer promotion eligible females across 600 plus “job fills” is *not* statistically significant by any measure, and does not support an inference of discrimination.

At this point the parties have merely presented two different statistical models that produce opposite results. Statistics “come in infinite variety and ... their usefulness depends on all of the surrounding facts and circumstances.” *Teamsters*, 431 U.S. at 341. Simply demonstrating that an alternative analysis leads to alternative results is not sufficient to defeat a plaintiff’s *prima facie* case - the defendant must also show that there is no genuine issue of material fact that plaintiffs’ model is fundamentally flawed for the purpose of demonstrating disparate impact in the case at issue. *Id.* (“[S]tatistics are not irrefutable ... like any other kind of evidence, they may be rebutted.”).

The key difference between the experts can be boiled down to this: Dr. Griffin looks at whether women applying to job X are disadvantaged relative to men applying to job X, whereas Dr. Killingsworth analyzes whether women applying to jobs X and Y are disadvantaged relative to men applying for jobs X and Y, combined. When seen in those terms, the difference between the expert analysis presented in this case is simply a question of whether the plaintiffs have analyzed the appropriate relevant labor pool for purposes of comparison. This question can be decided as a matter of law. *See e.g., Foxworth*, 228 Fed. App’x at 156 (holding in disparate impact case that where plaintiff’s statistics showed a disparity among police force staff, but not cadet applicants, they were insufficient as a matter of law to prove disparate impact as to cadet

“feeder job.” None of these groups contain multiple observations for the same employee. (Griffin Rep. 6.)

applicants).

2. *Plaintiffs' Model Does Not Reflect the Relevant Labor Pool in this Case*

In aggregating the feeder pool data, Dr. Killingsworth defines the relevant comparison as all union employees at Amtrak who are minimally qualified for a promotion to management, based on promotions that took place between March 8, 2002, and Jun 30, 2007. Plaintiffs argue that because the one-year rule applies to all union employees uniformly, it is appropriate to analyze the entire universe of feeder jobs uniformly to determine whether the blocking rule disqualifies women at a greater rate than men among this group as a whole.

Aggregated statistical data may be properly used to prove disparate impact where it is more probative than subdivided data. *Paige v. California*, 291 F.3d 1141, 1148 (9th Cir. 2002) (permitting plaintiffs to aggregate data from various supervisory written examinations where there was sufficient commonality among the duties and skills required by the different positions to justify aggregation); *see also Hazelwood Sch. Dist. v. U.S.*, 433 U.S. 299, 308 n.13 (1977) (aggregating various types of public school teacher positions for purposes of analysis where there were no special qualifications to consider). However, “[w]hen special qualifications are required to fill particular jobs, comparisons to the general population (rather than to the smaller group of individuals who possess the necessary qualifications) may have little probative value.” *Hazelwood Sch. Dist. v. U.S.*, 433 U.S. at 308 n.13; *Smithers v. Bailar*, 629 F.2d 892 (3d Cir. 1980) (holding that statistics in a disparate impact case “must bear some direct relationship to the applicant pool and the position sought”).

The evidence in the record regarding Amtrak’s hiring practices demonstrates that while Amtrak does not limit the positions for which existing employees may apply, minimal

qualifications are considered for every management position, and these qualifications vary from position to position.¹⁷ (See Pls.’ Mem. in Opp. to Def.’s Mot. for Summ. J. Exs. 3, 4, 12; Def.’s Reply In Support of Mot. Summ. J. 14 n.10; *see also* Ramirez Dep. 77:17-78:6, 204:15-205:19, 281:17-7, 335:2-6.) Dr. Killingsworth adopts this position, and in initially devising the Feeder Pools he assumes that the relevant labor market for each “job fill” is other employees who are presumptively qualified for the same “job fill.” Dr. Killingsworth does not assume that every union employee is fungible for the purposes of promotion, otherwise he would have simply compared all union employees across the board. (See Killingsworth Second Rep. ¶ 24.) If these distinctions between job categories are important (as Dr. Killingsworth suggests by painstakingly creating the Feeder Pools in the first place), then the defendant’s argument that these distinctions should be maintained throughout the analysis rings true. Combining the Feeder Pools together for the purposes of statistical analysis obfuscates the reality that a woman in Feeder Pool “A” may not be impacted by the one-year rule relative to the men in that Feeder Pool, while the opposite could be true for a woman in Feeder Pool “B.” While it is true that the one-year rule applies equally to all union employees, across feeder jobs, clearly there are some feeder jobs or “job fill” situations for which the one-year has little or no effect on women relative to similarly situated men. The single aggregated statistic Dr. Killingsworth relies on compares individuals who may never actually be in competition for the same jobs, and does not accurately account for what job the employee in question is coming from, where they are looking to go, and what the

¹⁷ The plaintiffs argue that Amtrak did not provide them with any guidance with respect to minimum qualifications for any management position, it does not follow from this that there *are* no minimum qualifications for management positions. The plaintiffs clearly recognize this fact as the rationale behind Dr. Killingsworth’s Feeder Pool concept implies.

relevant qualifications are. These are essential factors to evaluate in a disparate impact analysis. Statistical comparisons, “if they are to have any value, must be between comparable groups and free from variables which would undermine the reasonableness of discrimination inferences to be drawn.” *Mazus v. Dep’t of Transp., Commonwealth of Pa.*, 629 F.2d 870, 875 (3d Cir. 1980) (rejecting plaintiff’s census data for the job category “laborers except farm” where the types of workers it included were not comparable to highway maintenance workers, the job category at issue). Where there are clearly minimum qualifications for particular positions, only applicants possessing these qualifications should be included in the analysis. *See e.g., Mayor of City of Philadelphia v. Educational Equality League*, 415 U.S. 605, 620 (1974) (positing that in a case where the jobs at issue were restricted to the highest-ranking officers of designated categories of citywide organizations, “the relevant universe for comparison purposes consists of the highest ranking officers of the categories of organizations and institutions specified in the charter, not the population at large”); *compare Allen v. Seidman*, 881 F.2d 375, 379 (7th Cir. 1989) (Posner, J.) (permitting aggregated statistics comparing test scores of black and white bank exam test-takers across multiple job categories where there was evidence that the pool was “reasonably homogeneous despite possible differences...in original entry qualifications”). While “the population selected for statistical analysis need not perfectly match the pool of qualified persons,” without “a close fit between the population used to measure disparate impact and the population of those qualified for a benefit, the statistical results cannot be persuasive.” *Carpenter v. Boeing Co.*, 456 F.3d 1183, 1196 (10th Cir. 2006) (rejecting statistical analysis where the variables considered by the plaintiff’s expert in analyzing the effect of gender on overtime assignments did not correlate with the variables actually used by the employer in

assigning overtime to affected workers); *Mazus*, 629 F.2d at 875 (affirming district court’s finding that the plaintiff’s “statistical source did not accurately reflect the percentage of females interested in the work force in question, and thus did not establish a *prima facie* case”).

In defense of aggregating the Feeder Pools, the plaintiffs submit the affidavit of Ramona Paetzold, a professor of management at Texas A & M University as a rebuttal expert in statistics.¹⁸ Ms. Paetzold argues that stratification is inappropriate in this case because the numbers of women in each feeder job at any given point in time is determined, in part, by the existence of the one-year rule itself, “because the one-year rule at least partially affects how long men and women must remain in the feeder job before being eligible for promotion.” (Paetzold Aff. 3.) Paetzold further argues that because the Policy determines the gender composition of feeder jobs at points in time, neither feeder job nor time should be used as a statistical stratification variable for purposes of determining the effect of the Policy on women. Likewise, Dr. Killingsworth argues that controlling for differences in ineligibility rates in different feeder jobs, as Dr. Griffin does, “erase[s] the relation between female sex and ineligibility . . . because differences in ineligibility rates across feeder jobs (which in this case are an artifact of the one-year rule) will be likely to be correlated with the representation of women in those jobs.” (Killingsworth Third Aff. 13.) These arguments are not persuasive.

The plaintiffs have provided no evidence that could reasonably lead a jury to conclude that the one-year rule determines the gender composition of feeder jobs. True, the one-year rule dictates who in each Feeder Pool will be blocked from promotion at any given point in time, and

¹⁸ Paetzold is the co-author, with Steven Willborn, of *The Statistics of Discrimination: Using Statistical Evidence in Discrimination Cases* (2002), which Dr. Griffin cites as support for his stratification method.

this composition may change absent the one-year rule, but the gender composition of feeder jobs may very well be affected by additional factors such as wage levels, working conditions, movement prospects, layoffs, and the union's collectible bargaining agreement that allows unrestricted lateral job movements among union employees, none of which the plaintiffs have made any attempt to identify or control for in their analysis.¹⁹ Even if plaintiffs' argument that the gender makeup of feeder jobs would change absent the one-year rule were accepted as true, the plaintiffs have failed to provide any evidence to indicate how it would change, or that it would change in a way that would successfully rebut Dr. Griffin's analysis or conclusions.

This fundamental problem with plaintiffs' statistical evidence is not cured by Dr. Killingsworth's additional calculations which he performs in his Third Affidavit. In response to Dr. Griffin's early criticism of Dr. Killingsworth's Feeder Pool concept, Dr. Killingsworth adopts a broader definition of "feeder job" in his Third Affidavit: for every promotion to management he includes in the Feeder Pool incumbent "feeder job" data for every union position that ever acted as a feeder to the management position at issue. For example, if an individual in union job X was appointed in 2003 to Manager, and an individual in union job Y was appointed to Manager in 2005, then both union job X and union job Y would be treated as a feeder job for each of these two appointments to Manager. Using this broader data set, Dr. Killingsworth analyzes the data using the stratification method Dr. Griffin employs in his analysis, and finds that women would be deemed ineligible (because of the one-year rule) 1,000 - 1,600 times (7.890

¹⁹ In his original report, Dr. Killingsworth explicitly states that "[t]he studies discussed in this report are not concerned with the reasons why women might have less within-position service than men. Rather, these studies are concerned with whether the application of the one-year rule will adversely affect women relative to men." (Killingsworth Rep. footnote 5.)

- 11.978 standard deviations) more than would be expected if women and men had an equal ineligibility rates.

Dr. Killingsworth's new data set, which departs substantially from his original "feeder job" configuration, does not accurately encompass who may realistically be considered minimally qualified for and interested in a given promotion opportunity at Amtrak. In addition, Dr. Killingsworth's new data set is vulnerable to the criticism that it gives too much weight to certain individuals' eligibility status, leading to skewed results. Dr. Griffin illustrates the latter point in his Supplemental Report, in which he compares the data set from Dr. Killingsworth's original report (the "Original Killingsworth Data Set") with his revised data set (the "New Killingsworth Data Set") for those union jobs acting as feeders to the Road Foreman position.

For the period in question, there were a total of seventy (70) job fills from various union positions to the management title of Road Foreman. Forty-four (44) of those promotions came from the ranks of only two union positions, Off-Corr Engineer and Passenger Engineer. The other 26 promotions came from 14 other union jobs, most of which only appeared once in the Original Killingsworth Data Set, in keeping with their frequency of use as a source for promotion. Dr. Griffin calls attention to two of these one-time-only feeder positions: Assistant Passenger Conductor/Trainee and Assignment Clerk. These two union jobs acted as a feeder for a promotion to Road Foreman only one time each, but contributed a combined 9 ineligible women to the sum total of 18 ineligible females in the Original Killingsworth Data Set (summed over all 70 job fills). By contrast, in the New Killingsworth Data Set, Dr. Killingsworth includes the incumbent data for the Assistant Passenger Conductor/Trainee and Assignment Clerk jobs for each one of the 70 promotion events. The total surplus of ineligible females resulting from these

two jobs alone becomes 213, which averages to 203 ineligible females when summed across all 70 job fills for the Road Foreman position. When looking at the revised data, it becomes clear that this ten-fold increase in the net surplus calculation of ineligible females is principally attributable to the Assistant Passenger/Trainee and Assignment Clerk jobs. These positions, which only acted as a feeder job for the Road Foreman position once, now account for the majority of ineligible women in the pool.

If anything, plaintiffs' exercise here supports Amtrak's argument that the impact of the one-year rule changes depending on which employees are assumed eligible for which jobs, and therefore, the proper analysis is a stratified one which compares employees against its peers within each relevant labor pool, and then aggregates the individual results. The relevant labor pool from which to analyze whether the one-year rule had a disparate impact on female union employees, is against similarly qualified and interested male union employees, which in this case is reasonably approximated as those individuals holding the same (or equivalent) jobs. The Feeder Pools that Dr. Killingsworth originally constructed appropriately reflected these categories, but because he aggregated the Feeder Pools before analyzing the impact of the one-year rule, the resultant statistics have insufficient probative value as evidence that one-year rule has a disparate impact on Amtrak's female union employees. In other words, because plaintiffs' analysis is focused on an overbroad and incomparable pool of employees, it lacks the statistical significance necessary to make out a *prima facie* case of discrimination.

b. Practical Significance

Even if Dr. Killingsworth's methodology was sound and his results recognized as having "statistical significance," the results of his analysis are undermined by a lack of practical

significance. “Statistical significance” is a term of art within the science of statistics which means simply that the disparity was unlikely to have been produced by chance. Statistical significance is routinely established in cases where observations are made among a reasonably large population. *See Carpenter*, 456 F.3d at 1201 (“[S]tatistical significance does not necessarily mean that the departure from equality was large.”). For this reason, courts look for practical significance as an important signifier of impact in addition to statistical significance. *See* Barbara Lindemann & Paul Grossman, *Employment Discrimination Law* 94 (3d ed. 1996) (“To guard against the possibility that a finding of adverse impact could result from the statistical significance of a trivial disparity or from a meaningless difference in results, the Uniform Guidelines on Employment Selection Procedures, and the courts have adopted an additional test for adverse impact: that a statistically significant disparity also has practical significance.”); *see also* 29 C.F.R. § 1607.4D.

In evaluating the practical significance of Dr. Killingsworth’s results, Dr. Griffin calls attention to the adverse impact ratio Dr. Killingsworth found, which was 96.8 percent (80% being the figure at or below which the EEOC will presume the existence of adverse impact). This rate is statistically significant at conventional test levels, but its practical significance is of limited magnitude. Based on the numbers used by Dr. Killingsworth in his first report, Dr. Griffin calculated that if female candidates in the Aggregated Feeder Pool (of which there were 29,919) had the same eligibility rate as males, this would translate to a “gender gap” of only 726 additional female promotion-eligible situations overall.²⁰ This means that the eligibility rates by

²⁰ This rate is found by multiplying the pass rate difference between men and women of 2.42% (= 74.65% - 72.23%) by the total number of female candidacies in the Aggregated Feeder Pool (29,919). (Griffin Rep. 5.)

gender would have been identical for men and women if roughly one more female had met the time-in-position rule for each of the promotion events that Dr. Killingsworth considered.²¹ As a consequence of that one additional eligible female, the proportion of “eligible” females in the Aggregated Feeder Pool would have increased from approximately 22.08 % to 22.7%.²²

In addition, Dr. Griffin submits that even with the one-year rule in effect, females received a greater than expected share of the actual promotions being looked at. Of the promotion events examined by Dr. Killingsworth, females received 23.56% of the promotions. However, only 22.08% of females in the Aggregated Feeder Pool are promotion eligible, indicating that females received a larger percentage of promotions than their eligibility rates would otherwise suggest.

These negligible figures, in addition to the adverse impact ratio of 96.8%, are not persuasive of a finding of disparate impact, statistical significance notwithstanding. *See e.g., Waisome v. Port Authority of New York and New Jersey*, 948 F.2d 1370 (2d Cir. 1991) (affirming district court’s finding of no substantial significance where the disparity was “statistically significant, [but] it was of limited magnitude” in light of the fact that if two additional black

²¹ It is important to remember that the 726 number does not represent actual female employees who were actually rejected for promotions. This figure represents the number of times that an additional female union employee with more than one-year of service would have needed to appear in the Aggregated Feeder Pool to achieve “gender parity” in ineligibility rates for men and women in the pool (“gender parity” meaning that male and female employees would have a numerically identical likelihood of having less than one year of time-in-position). So this “gender gap” figure may translate to one female for each of 726 different promotions, or two females each for 363 different positions, or some other configuration.

²² The 22.08% figure represents the 21,608 eligible women out of a total of 97,869 eligible employees in the Aggregated Feeder Pool ($21,608 / 97,869 = 22.08$) (Griffin Rep. 3; Killingsworth Rep. Table 1.)

candidates had passed the written exam at issue, the statistical disparity in pass rates between whites and blacks would have disappeared). The plaintiffs have failed to demonstrate that their evidence, when viewed in context, has the degree of practical significance necessary to raise an inference of discrimination.

V. Conclusion

The applicant pool plaintiffs analyzed to demonstrate the disparate impact of Amtrak's policy erroneously compares employees who may not have the minimal qualifications for the particular jobs at issue. Furthermore, when viewed in context, plaintiffs' evidence of discrimination lacks practical significance. The plaintiffs have therefore failed to carry their burden of presenting a *prima facie* case of disparate impact discrimination under Title VII. Accordingly, summary judgment is granted.

s/Anita B. Brody

ANITA B. BRODY, J.

**IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF PENNSYLVANIA**

SHARYN STAGI and WINIFRED LADD	:	CIVIL ACTION
Individually, and on behalf of	:	
all others similarly situated	:	
	:	
Plaintiffs,	:	
	:	NO. 03-CV-5702
v.	:	
	:	
NATIONAL RAILROAD PASSENGER	:	
CORPORATION (t/d/b/a AMTRAK)	:	
	:	
Defendant.	:	

ORDER

AND NOW, this ____12th ____ day of August 2009, it is **ORDERED** that Defendant's Motion for Summary Judgment (Doc. #77) is **GRANTED**.

s/Anita B. Brody

ANITA B. BRODY, J.